

Broadcast Channels with Confidential Messages by Randomness Constrained Stochastic Encoder

Shun Watanabe *Member, IEEE* and Yasutada Oohama *Member, IEEE*

Abstract—In coding schemes for the wire-tap channel or the broadcast channels with confidential messages, it is well known that the sender needs to use a stochastic encoding to avoid the information about the transmitted confidential message to be leaked to an eavesdropper. In this paper, it is investigated that the trade-off between the rate of the random number to realize the stochastic encoding and the rates of the common, private, and confidential messages. For the direct theorem, the superposition coding scheme for the wire-tap channel recently proposed by Chia and El Gamal is employed, and its strong security is proved. The matching converse theorem is also established. Our result clarifies that a combination of the ordinary stochastic encoding and the channel prefixing by the channel simulation is suboptimal.

Index Terms—Broadcast Channel, Confidential Messages, Randomness Constraint, Stochastic Encoder, Superposition Coding, Wire-tap Channel

I. INTRODUCTION

The wire-tap channel is one sender and two receivers broadcast channel model in which the sender, usually referred to as Alice, wants to transmit a confidential message to the legitimate receiver, usually referred to as Bob, in such a way that the other receiver, usually referred to as eavesdropper Eve, cannot get any information about the transmitted message. The wire-tap channel model was first introduced by Wyner in his seminal paper [1]. Later, Csiszár and Körner investigated the model called broadcast channels with confidential messages (BCC) in which Alice also sends a common message that is supposed to be decoded by both Bob and Eve. These models were further investigated by many researchers from theoretical point of view (e.g., see [2]), and recently it has attracted considerable attention from practical point of view as a physical layer security.

In coding schemes for the wire-tap channel or the BCC, it is well known that the sender needs to use a stochastic encoder to avoid the information about the transmitted confidential message to be leaked to Eve. The stochastic encoding is usually realized by preparing a dummy random number in addition to the intended messages and by encoding them to a transmitted signal by a deterministic encoder. Furthermore, when the channel to Bob is not more capable than the channel to Eve, it is known that the sender needs to use the channel prefixing to achieve the capacity region (or the secrecy capacity) because the capacity formulas involve

such a channel from an auxiliary random variable to the random variable describing the input signal of the channel [3]. In literatures, it is assumed that there exists a channel realizing the channel prefixing. But in practice the prefixing channel must be simulated from a random number by using a method such as the channel simulation [4], which usually involves certain amount of simulation error depending on the amount of the random number. So far, there was no paper investigating how much random number is needed to achieve the capacity region. Since the random number is precious resource in practice, though it has been paid no attention in literatures, it is extremely important to investigate the amount of random number needed to achieve the capacity region. For this purpose, we formulate the problem of the BCC by randomness constrained stochastic encoder, and completely characterize the capacity region of this new problem.

The present problem to consider the randomness constrained stochastic encoder is motivated by the authors' previous results in [5]. In that paper, the capacity region of the relay channel with confidential messages for the completely deterministic encoder was investigated, and the capacity region was characterized for the BCC as a corollary. In this paper, we are interested in the case such that the randomness is constrained but not zero. The result in [5] can be regarded as an extreme case of the present problem. On the other hand, the conventional BCC problem can be regarded as the other extreme case, in which the amount of randomness that can be used at the encoder is unbounded.

Typically in the BCC, Alice sends the common message that is supposed to be decoded by both Bob and Eve, and the confidential message that is supposed to be decoded only by Bob. The level of secrecy of the confidential message is usually evaluated by the equivocation rate. In this paper, we consider slightly different problem formulation, which has been appeared in the literature [6], [5]. In our problem setting, Alice sends three kinds of messages, the common message, the private message, and the confidential messages. The common message is supposed to be decoded by both Bob and Eve. The private message is supposed to be decoded by Bob, and we do not care whether Eve can decode the private message or not. The confidential message is supposed to be decoded by Bob, and it must be kept completely secret from Eve. Furthermore, for stochastic encoding, Alice is allowed to use limited amount of dummy randomness. Thus, we are interested in the trade-off between quadruple of rates, the rate of dummy randomness, the rates of common, private, and confidential messages. The coding system of our formulation is depicted in Fig. 1.

The reason we do not use the equivocation rate formulation is as follows. In the conventional equivocation rate formula-

The first author is with the Department of Information Science and Intelligent Systems, University of Tokushima, 2-1, Minami-josanjima, Tokushima, 770-8506, Japan, e-mail:shun-wata@is.tokushima-u.ac.jp.

The second author is with the Department of Communication Engineering and Informatics, University of Electro-Communications, Tokyo, 182-8585, Japan, e-mail:oohama@uec.ac.jp.

Manuscript received ; revised

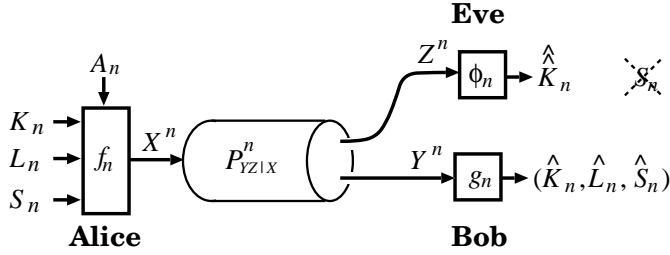


Fig. 1. The coding system investigated in this paper. Alice sends common message K_n , private message L_n , and confidential message S_n by using a deterministic function f_n and a limited amount of dummy randomness A_n . The common message is supposed to be decoded by both Bob and Eve. The private message is supposed to be decoded by Bob, and we do not care whether Eve can decode the private message or not. The confidential message is supposed to be decoded by Bob, and it must be kept completely secret from Eve.

tion, if the rate of dummy randomness is not sufficient, a part of the confidential message is sacrificed to make the other part completely secret and the rate of the completely secret part corresponds to the equivocation rate. We think that the rates of sacrificed part and completely secret part become clearer by employing our formulation.

As we have mentioned above, the typical coding scheme for the wire-tap channel or the BCC consists of the stochastic encoding and the channel prefixing. In [7], Chia and El Gamal proposed an alternative coding scheme that utilizes the so-called superposition coding [8] instead of the channel prefixing. In the direct part of our main result, we employ this superposition scheme instead of the channel prefixing. We also clarify that a straightforward combination of the ordinary stochastic coding and the channel prefixing by the channel simulation method is suboptimal.

Although Chia and El Gamal showed that the superposition coding scheme can realize the so-called weak security criterion [9], [10], it was not clear whether the superposition coding scheme can realize the so-called strong security criterion. One of technical contributions of this paper is to show that Chia and El Gamal's superposition coding scheme can realize the strong security criterion. This is done by using the technique proposed in [11], and by considering the channel resolvability problem [12] with the superposition coding. Note that the relationship between the wire-tap channel coding and the channel resolvability was first pointed out by Csiszár [10], and is well recognized recently [13], [14], [15], [16]. The channel resolvability with the superposition coding was first investigated by the second author in [17]. In that paper, the channel resolvability problem with the superposition coding for the degraded broadcast channel was considered to show the converse theorem of the identification via degraded broadcast channels. In this paper, the channel resolvability problem with the superposition coding for a single channel is considered. Using the superposition coding for a single channel seems nonsense at first glance, it does have a meaning when applied to the wire-tap channel or the BCC.

After the submission of the first manuscript of this paper, we noticed some related works investigating the importance

of random number in the BCC or the wire-tap channel. In [18], Hayashi and Matsumoto considered the secure multiplex coding [19] in which the messages are not necessarily uniform nor independent and the entropy rate of the messages might be constrained. Although the secure multiplex coding can be regarded as a generalization of the BCC, the encoder can use unlimited amount of dummy randomness in addition to the messages in their problem formulation. Thus, our results cannot be derived from their results. In [20], Bloch and Klierer considered the wire-tap channel in which the dummy randomness is constrained and not necessarily uniform. However, they only considered the case such that the channel to Bob is more capable than that to Eve. In such a case, the channel prefixing is not needed, and their result corresponds to Corollary 12 in this paper when the dummy randomness is uniform.

The rest of the paper is organized as follows. In Section II, the problem formulation is explained and main results are presented. In Section III, the channel resolvability problem with the superposition coding is discussed. In Section IV, the proof of the main theorem is presented. In Section V, the paper is concluded with discussions. Some technical arguments are presented in Appendices.

II. PROBLEM FORMULATION AND MAIN RESULTS

A. Problem Formulation

Let $P_{Y|X}$ and $P_{Z|X}$ be two channels with common input alphabet \mathcal{X} and output alphabets \mathcal{Y} and \mathcal{Z} respectively. Throughout the paper, the alphabets are assumed to be finite though we do not use finiteness of the alphabet except cardinality bounds on auxiliary random variables. We also assume that the base of logarithm is e throughout the paper.

Let \mathcal{K}_n be the set of the common message, \mathcal{L}_n be the set of the private message, and \mathcal{S}_n be the set of the confidential message. The common message is supposed to be decoded by both Bob and Eve. The private message is supposed to be decoded by Bob, and we do not care whether Eve can decode the private message or not. The confidential message is supposed to be decoded by Bob, and it must be kept completely secret from Eve.

Typically, Alice use a stochastic encoder to make the confidential message secret from Eve, and it is practically realized by using a uniform dummy randomness on the alphabet \mathcal{A}_n . When the size $|\mathcal{A}_n|$ of dummy randomness is infinite, any stochastic encoder from $\mathcal{K}_n \times \mathcal{L}_n \times \mathcal{S}_n$ to \mathcal{X}^n can be simulated by a deterministic encoder $f_n : \mathcal{K}_n \times \mathcal{L}_n \times \mathcal{S}_n \times \mathcal{A}_n \rightarrow \mathcal{X}^n$. But we are interested in the case with bounded size $|\mathcal{A}_n|$ in this paper.

Bob's decoder is defined by function $g_n : \mathcal{Y}^n \rightarrow \mathcal{K}_n \times \mathcal{L}_n \times \mathcal{S}_n$ and the error probability is defined as

$$\begin{aligned}
 P_{err}(f_n, g_n) &= \sum_{k_n \in \mathcal{K}_n} \sum_{\ell_n \in \mathcal{L}_n} \sum_{s_n \in \mathcal{S}_n} \sum_{a_n \in \mathcal{A}_n} \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n| |\mathcal{S}_n| |\mathcal{A}_n|} \\
 &\quad P_{Y|X}^n(y^n | f_n(k_n, \ell_n, s_n, a_n)) \mathbf{1}[g_n(y^n) \neq (k_n, \ell_n, s_n)],
 \end{aligned} \tag{1}$$

where $\mathbf{1}[\cdot]$ is the indicator function. Eve's decoder is defined by function $\phi_n : \mathcal{Z}^n \rightarrow \mathcal{K}_n$ and the error probability $P_{err}(f_n, \phi_n)$ is defined in a similar manner as Eq. (1).

Let

$$\begin{aligned} P_{\tilde{Z}^n|S_n}(z^n|s_n) &= \sum_{k_n \in \mathcal{K}_n} \sum_{\ell_n \in \mathcal{L}_n} \sum_{a_n \in \mathcal{A}_n} \frac{1}{|\mathcal{K}_n||\mathcal{L}_n||\mathcal{A}_n|} \\ &\quad P_{Z|X}^n(z^n|f_n(k_n, \ell_n, s_n, a_n)), \\ P_{\tilde{Z}^n}(z^n) &= \sum_{s_n \in \mathcal{S}_n} \frac{1}{|\mathcal{S}_n|} P_{\tilde{Z}^n|S_n}(z^n|s_n) \end{aligned}$$

be the output distributions of the channel $P_{Z|X}^n$. In this paper, we consider the security criterion given by

$$\begin{aligned} D(f_n) &:= D(P_{S_n \tilde{Z}^n} \| P_{S_n} \times P_{\tilde{Z}^n}) \\ &= \sum_{s_n \in \mathcal{S}_n} \frac{1}{|\mathcal{S}_n|} D(P_{\tilde{Z}^n|S_n}(\cdot|s_n) \| P_{\tilde{Z}^n}) \\ &= I(S_n; \tilde{Z}^n), \end{aligned}$$

where $D(\cdot \| \cdot)$ is the divergence, and $I(\cdot; \cdot)$ is the mutual information [8]. The coding system investigated in this paper is depicted in Fig. 1.

In this paper, we are interested in the trade-off among the rate of the dummy randomness, and the rates of the common, private, and confidential messages.

Definition 1: The rate quadruple (R_d, R_0, R_1, R_s) is said to be *achievable* if there exists a sequence of Alice's deterministic encoder $f_n : \mathcal{K}_n \times \mathcal{L}_n \times \mathcal{S}_n \times \mathcal{A}_n \rightarrow \mathcal{X}^n$, Bob's decoder $g_n : \mathcal{Y}^n \rightarrow \mathcal{K}_n \times \mathcal{L}_n \times \mathcal{S}_n$, and Eve's decoder $\phi_n : \mathcal{Z}^n \rightarrow \mathcal{K}_n$ such that

$$\lim_{n \rightarrow \infty} P_{err}(f_n, g_n) = 0, \quad (2)$$

$$\lim_{n \rightarrow \infty} P_{err}(f_n, \phi_n) = 0, \quad (3)$$

$$\lim_{n \rightarrow \infty} D(f_n) = 0, \quad (4)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{A}_n| \leq R_d, \quad (5)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{K}_n| \geq R_0, \quad (6)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{L}_n| = R_1, \quad (7)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{S}_n| \geq R_s. \quad (8)$$

Then the achievable region \mathcal{R} is defined as the set of all achievable rate quadruples.

B. Statements of General Results

The following is our main result in this paper.

Theorem 2: Let \mathcal{R}^* be a closed convex set consisting of those quadruples (R_d, R_0, R_1, R_s) for which there exist auxiliary random variables (U, V) such that $U \leftrightarrow V \leftrightarrow X \leftrightarrow$

(Y, Z) and

$$R_0 \leq \min[I(U; Y), I(U; Z)], \quad (9)$$

$$R_0 + R_1 + R_s \leq I(V; Y|U) + \min[I(U; Y), I(U; Z)], \quad (10)$$

$$R_s \leq I(V; Y|U) - I(V; Z|U), \quad (11)$$

$$R_1 + R_d \geq I(X; Z|U), \quad (12)$$

$$R_d \geq I(X; Z|V). \quad (13)$$

Then we have $\mathcal{R} = \mathcal{R}^*$. Moreover, it may be assumed that $V = (U, V')$ and that the ranges of U and V' may be assumed to satisfy $|\mathcal{U}| \leq |\mathcal{X}| + 3$ and $|\mathcal{V}'| \leq |\mathcal{X}| + 1$.

Proof: See Section IV. ■

The conditions on R_0 and $R_1 + R_s$ in Eqs. (9) and (10) resemble the conditions in the broadcast channel with degraded message sets [21]. The condition on R_s in Eq. (11) exists because there is a security requirement on the confidential message. These conditions are exactly the same as those in the conventional BCC (see Corollary 6). The conditions on R_1 and R_d in Eqs. (12) and (13) additionally appear in Theorem 2 because there are randomness constraints in our problem setting.

Remark 3: Conventionally, the security requirement defined by

$$\lim_{n \rightarrow \infty} \frac{1}{n} D(f_n) = 0 \quad (14)$$

is usually employed instead of Eq. (4). Eq. (14) is called weak security criterion and Eq. (4) is called strong security criterion [10], [9]. Let $\tilde{\mathcal{R}}$ be the achievable region in which Eq. (4) is replaced by Eq. (14). From the definitions of two regions, $\mathcal{R} \subset \tilde{\mathcal{R}}$ obviously holds. Actually, we are implicitly showing $\tilde{\mathcal{R}} \subset \mathcal{R}^*$ in the converse proof of Theorem 2. Thus, $\mathcal{R} = \tilde{\mathcal{R}}$.

Remark 4: As we will find in the achievability proof of the main theorem, the private message can be used as dummy randomness to protect the confidential message from Eve. Thus, if we define the achievable rate region $\tilde{\mathcal{R}}$ by replacing Eq. (7) with

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{L}_n| \geq R_1, \quad (15)$$

region $\hat{\mathcal{R}}$ is broader than region \mathcal{R} . Indeed, $\hat{\mathcal{R}}$ is a closed convex set consisting of those quadruples (R_d, R_0, R_1, R_s) for which there exist auxiliary random variables (U, V) satisfying the same conditions as Theorem 2 except Eq. (12)¹.

Remark 5: Eq. (13) means that there is a certain amount of dummy randomness that cannot be substituted by the private message. Note that the difference between the private message and the dummy randomness is whether Bob needs to decode it or not.

Let

$$\mathcal{R}_\infty = \{(R_0, R_1, R_s) : \exists R_d \geq 0 \text{ s.t. } (R_d, R_0, R_1, R_s) \in \mathcal{R}\}$$

¹It can be proved by just omitting the derivation of Eq. (12) in the converse proof of Theorem 2.

be the set of all achievable triplet (R_0, R_1, R_s) by arbitrary stochastic encoder. By taking sufficiently large R_d , we recover the following well known result [3]².

Corollary 6: ([3]) Region \mathcal{R}_∞ is a closed convex set consisting of those triplet (R_0, R_1, R_s) for which there exist auxiliary random variables (U, V) such that $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$\begin{aligned} R_0 &\leq \min[I(U; Y), I(U; Z)], \\ R_0 + R_1 + R_s &\leq I(V; Y|U) + \min[I(U; Y), I(U; Z)], \\ R_s &\leq I(V; Y|U) - I(V; Z|U). \end{aligned}$$

Let

$$\mathcal{R}_{det} = \{(R_0, R_1, R_s) : (0, R_0, R_1, R_s) \in \mathcal{R}\}$$

be the set of all rate triplets that can be achieved by deterministic encoder. This extreme case was solved in [5], which can be also derived as a corollary of Theorem 2³.

Corollary 7: ([5]) Let \mathcal{R}_{det}^* be a closed convex set consisting of those triplet (R_0, R_1, R_s) for which there exists an auxiliary random variable U such that $U \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$\begin{aligned} R_0 &\leq \min[I(U; Y), I(U; Z)], \\ R_0 + R_1 + R_s &\leq I(X; Y|U) + \min[I(U; Y), I(U; Z)], \\ R_s &\leq I(X; Y|U) - I(X; Z|U), \\ R_1 &\geq I(X; Z|U). \end{aligned}$$

Then we have $\mathcal{R}_{det} = \mathcal{R}_{det}^*$.

Proof: The inclusion $\mathcal{R}_{det}^* \subset \mathcal{R}_{det}$ is obvious by taking $V = X$ in Theorem 2. For the opposite inclusion, note that Eq. (13) and $R_d = 0$ imply

$$I(X; Z|U) = I(V; Z|U) + I(X; Z|V) = I(V; Z|U).$$

We also have $I(V; Y|U) \leq I(X; Y|U)$ from the Markov condition of the auxiliary random variables. Thus, we have

$$\mathcal{R}_{det} \subset \{(R_0, R_1, R_s) : (0, R_0, R_1, R_s) \in \mathcal{R}^*\} \subset \mathcal{R}_{det}^*.$$

Let

$$R_d(R_0, R_s) = \inf\{R_d : (R_d, R_0, 0, R_s) \in \mathcal{R}\}$$

be the infimum rate of dummy randomness needed to achieve the rates (R_0, R_s) . From Theorem 2, we can characterize not only the known extreme cases (Corollary 6 and Corollary 7) but also this quantity.

Corollary 8: $R_d(R_0, R_s)$ is the optimal solution of the following optimization problem:

$$\begin{aligned} &\text{minimize} && I(X; Z|U) \\ &\text{subject to} && \\ &R_0 &\leq & \min[I(U; Y), I(U; Z)], \\ &R_0 + R_s &\leq & I(V; Y|U) + \min[I(U; Y), I(U; Z)], \\ &R_s &\leq & I(V; Y|U) - I(V; Z|U), \end{aligned}$$

²See also [6, Theorem 17.13] for the result that does not employ the rate-equivocation formulation.

³In [5], slightly deferent problem formulation is employed and the achievable region seems slightly different from Corollary 7. But they are essentially the same.

where (U, V) satisfy $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$.

Let

$$\mathcal{R}_{ds} = \{(R_d, R_s) : (R_d, 0, 0, R_s) \in \mathcal{R}\}.$$

As a corollary of Theorem 2, we also have the following.

Corollary 9: Let \mathcal{R}_{ds}^* be a closed convex set consisting of those rate pair (R_d, R_s) for which there exist auxiliary random variables (U, V) such that $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$R_s \leq I(V; Y|U) - I(V; Z|U), \quad (16)$$

$$R_d \geq I(X; Z|U). \quad (17)$$

Then we have $\mathcal{R}_{ds} = \mathcal{R}_{ds}^*$.

Remark 10: The auxiliary random variable U in Corollary 9 only plays a role of time-sharing. Thus, the range of U may be assumed to satisfy $|U| \leq 2$. The same remark is also applied for Corollary 12.

Let

$$C_s = \sup\{R_s : \exists R_d \geq 0 \text{ s.t. } (R_d, R_s) \in \mathcal{R}_{ds}\}$$

be the secrecy capacity, which can be characterized by the supremum of the rate R_s for which there exists auxiliary random variable V such that $V \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$R_s \leq I(V; Y) - I(V; Z). \quad (18)$$

To achieve the rate given by the right hand side of Eq. (18), we conventionally used the following coding scheme. First, we construct a wire-tap channel code for channel pairs $P_{Y|V}$ and $P_{Z|V}$. Then, the code word in \mathcal{V}^n is transmitted over prefixing channel $P_{X|V}^n$. If we simulate channel $P_{X|V}^n$ by using the channel simulation method [4], then we need randomness with rate $H(X|V)$ ⁴. By using this argument, we can derive the following inner bound on \mathcal{R}_{ds} that can be achieved by combining the ordinary wire-tap channel coding and the channel prefixing by the channel simulation method.

Proposition 11: Let \mathcal{R}_{sim}^* be a closed convex set consisting of those rate pair (R_d, R_s) for which there exist auxiliary random variables (U, V) such that $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$R_s \leq I(V; Y|U) - I(V; Z|U), \quad (19)$$

$$R_d \geq I(V; Z|U) + H(X|V). \quad (20)$$

Then we have $\mathcal{R}_{sim}^* \subset \mathcal{R}_{ds}$.

Since $I(X; Z|U) = I(V; Z|U) + I(X; Z|V) < I(V; Z|U) + H(X|V)$ in general, the region \mathcal{R}_{ds}^* is strictly broader than the region \mathcal{R}_{sim}^* , i.e., the straightforward combination of the ordinary wire-tap channel coding and the channel prefixing by the channel simulation is suboptimal.

Corollary 12: Suppose that the channel $P_{Y|X}$ is more capable than $P_{Z|X}$. Then the region $\mathcal{R}_{ds} = \mathcal{R}_{ds}^*$ is a closed convex set consisting of those rate pair (R_d, R_s) for which there exists an auxiliary random variable U such that $U \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$R_s \leq I(X; Y|U) - I(X; Z|U),$$

$$R_d \geq I(X; Z|U).$$

⁴We are implicitly assuming that the empirical distributions of almost every code words are close to P_V , which is true if we use the random coding method.

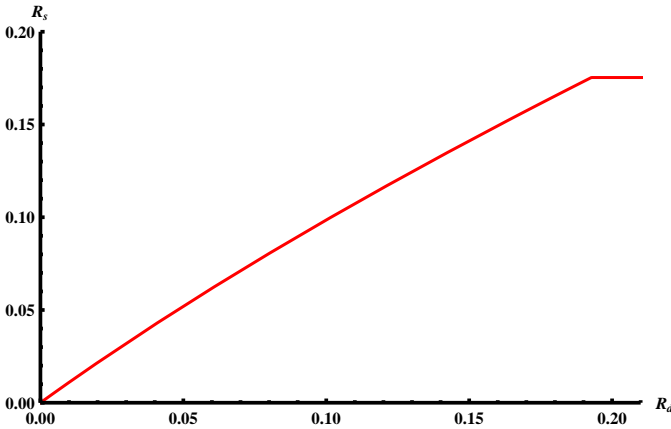


Fig. 2. The achievable region \mathcal{R}_{ds} when $P_{Y|X}$ is BSC(0.1) and $P_{Z|X}$ is BSC(0.2).

Moreover, it may be assumed that the ranges of U may be assumed to satisfy $|U| \leq 2$.

Proof: See Appendix E. ■

As we can find from Corollary 12, we do not need auxiliary random variable V when the channel $P_{Y|X}$ is more capable than $P_{Z|X}$. Thus, two regions \mathcal{R}_{ds}^* and \mathcal{R}_{sim}^* coincide.

C. Numerical Examples

First, we consider an example such that \mathcal{R}_{ds}^* and \mathcal{R}_{sim}^* coincide. Suppose that $P_{Y|X}$ and $P_{Z|X}$ are binary symmetric channels with crossover probabilities ε_1 and ε_2 respectively, where $\varepsilon_1 < \varepsilon_2$. In this case, $P_{Z|X}$ is degraded version of $P_{Y|X}$, which also implies that $P_{Y|X}$ is more capable than $P_{Z|X}$. Thus, we can apply Corollary 12. Since the auxiliary random variable U only plays a role of time sharing, region \mathcal{R}_{ds} is the convex hull of the rates (R_d, R_s) satisfying

$$\begin{aligned} R_s &\leq [h(p * \varepsilon_1) - h(\varepsilon_1)] - [h(p * \varepsilon_2) - h(\varepsilon_2)], \\ R_d &\geq h(p * \varepsilon_2) - h(\varepsilon_2) \end{aligned}$$

for some input distribution $0 \leq P_X(0) = p \leq 1$, where $h(\cdot)$ is the binary entropy function⁵ and $x * y = x(1 - y) + (1 - x)y$ is the binary convolution. In Fig. 2, for the case with $\varepsilon_1 = 0.1$ and $\varepsilon_2 = 0.2$ respectively, the region \mathcal{R}_{ds} is plotted. The input distribution achieving C_s is the uniform distribution, and thus R_s is constant when $R_d \geq \log 2 - h(0.2)$. By using a biased input distribution, R_s can be positive even if R_d is smaller than $\log 2 - h(0.2)$.

Next, we consider an example such that \mathcal{R}_{ds}^* and \mathcal{R}_{sim}^* do not coincide. Suppose that $P_{Y|X}$ is a binary symmetric channel with crossover probability ε and $P_{Z|X}$ is a binary erasure channel with erasure probability δ . When $4\varepsilon(1 - \varepsilon)\log 2 < \delta \log 2 < h(\varepsilon)$, it is known that $P_{Y|X}$ is not more capable than $P_{Z|X}$ [22] and $C_s > 0$ ⁶. For this example, we can compute the regions $\mathcal{R}_{ds} = \mathcal{R}_{ds}^*$ as follows. Since \mathcal{R}_{ds}^* is a convex set,

⁵Note that the base of the logarithm is e .

⁶Actually, for $4\varepsilon(1 - \varepsilon)\log 2 < \delta \log 2 < h(\varepsilon)$, $P_{Z|X}$ is more capable than $P_{Y|X}$ but $P_{Z|X}$ is not less noisy than $P_{Y|X}$ [22]. Thus, $I(X; Y) \leq I(X; Z)$ for every P_X but there exists V such that $I(V; Y) > I(V; Z)$, which means that $C_s > 0$ and V is needed to achieve C_s .

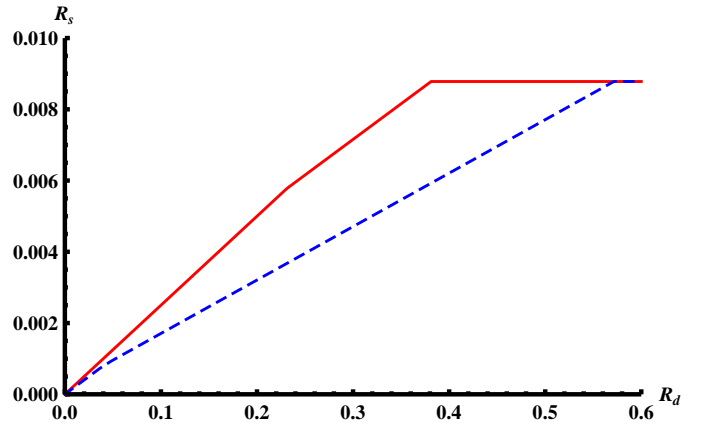


Fig. 3. The achievable region $\mathcal{R}_{ds} = \mathcal{R}_{ds}^*$ (solid line) and suboptimal inner bound \mathcal{R}_{sim}^* (dashed line) when $P_{Y|X}$ is BSC(0.11) and $P_{Z|X}$ is BEC(0.45).

for each R_d , we can calculate $\max\{R_s : (R_d, R_s) \in \mathcal{R}_{ds}^*\}$ by minimizing

$$\max_{P_{UVX}} [I(V; Y|U) - I(V; Z|U) - \mu(I(X; Z|U) - R_d)] \quad (21)$$

with respect to $\mu \geq 0$, where μ is the slope of the supporting line of \mathcal{R}_{ds}^* . Since U only plays the role of the times sharing in Eq. (21), we can take U to be constant. Furthermore, by using the support lemma [6], we can assume that $|\mathcal{V}| \leq |\mathcal{X}| = 2$. Thus, Eq. (21) can be calculated by exhaustive search of three parameters $P_V(0)$, $P_{X|V}(0|0)$, and $P_{X|V}(1|1)$. Since $P_V(0) = \frac{1}{2}$ is not necessarily optimal⁷ for $R_d < (1 - \delta)\log 2$, further reduction of parameters seems difficult. The region \mathcal{R}_{sim}^* can be computed in a similar manner.

In Fig. 3, for the case with $\varepsilon = 0.11$ and $\delta = 0.45$ respectively, the region $\mathcal{R}_{ds} = \mathcal{R}_{ds}^*$ and \mathcal{R}_{sim}^* are plotted.

III. CHANNEL RESOLVABILITY BY SUPERPOSITION

In this section, we consider the channel resolvability problem. The result in this section will be utilized in the direct part of the proof of Theorem 2.

In the channel resolvability problem, we want to simulate the response P_Z^n by using channel $P_{Z|X}^n$ and as small number of uniform randomness as possible, where P_Z^n is the n th product of

$$P_Z(z) = \sum_x P_X(x) P_{Z|X}(z|x)$$

for input distribution P_X . The simulation is conducted by a deterministic map $\varphi_n : \mathcal{B}_n \rightarrow \mathcal{X}^n$ and uniform random number B_n on \mathcal{B}_n . Let

$$P_{\tilde{Z}^n}(z^n) = \sum_{b_n \in \mathcal{B}_n} \frac{1}{|\mathcal{B}_n|} P_{Z|X}^n(z^n | \varphi_n(b_n))$$

be the output distribution with map φ_n . In this paper, the accuracy of the simulation is evaluated by the divergence

⁷When there is no constraint on R_d , it is known that $P_V(0) = \frac{1}{2}$ and $P_{X|V}(0|0) = P_{X|V}(1|1)$ are optimal [23].

criterion $D(P_{\tilde{Z}^n} \| P_Z^n)$. It is well known [12]⁸ that if

$$R > I(X; Z), \quad (22)$$

then there exists a sequence of maps $\{\varphi_n\}$ satisfying

$$\lim_{n \rightarrow \infty} D(P_{\tilde{Z}^n} \| P_Z^n) = 0, \quad (23)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n| \leq R. \quad (24)$$

Typically, a sequence of maps realizing Eqs. (23) and (24) is constructed by randomly generating $|\mathcal{B}_n|$ codeword $x_1^n, \dots, x_{|\mathcal{B}_n|}^n$ according to P_X^n . We denote the generated code \mathcal{C}_n . Then we have the following proposition.

Proposition 13: ([11]) For every $n \geq 1$, we have

$$\mathbb{E}_{\mathcal{C}_n} [D(P_{\tilde{Z}^n} \| P_Z^n)] \leq \frac{1}{\theta |\mathcal{B}_n|^\theta} e^{n\psi(\theta|P_{Z|X}, P_X)}$$

for $0 < \theta \leq 1$, where $\mathbb{E}_{\mathcal{C}_n}[\cdot]$ means taking the average over the randomly generated code \mathcal{C}_n , and the function $\psi(\theta|P_{Z|X}, P_X)$ is defined by

$$\begin{aligned} \psi(\theta|P_{Z|X}, P_X) &= \log \sum_z \left(\sum_x P_X(x) P_{Z|X}(z|x)^{1+\theta} \right) P_Z(z)^{-\theta}. \end{aligned} \quad (25)$$

In this paper we construct a sequence of maps realizing Eq. (23) by a different method. Let P_{VX} be a distribution such that the marginal is P_X . We first randomly generate $|\mathcal{M}_{2,n}|$ codeword $v_1^n, \dots, v_{|\mathcal{M}_{2,n}|}^n$ according to the distribution P_V^n . We denote the generated code by $\mathcal{C}_{2,n}$. Then for each $1 \leq i \leq |\mathcal{M}_{2,n}|$, we randomly generate $|\mathcal{M}_{1,n}|$ codeword $x_{i1}^n, \dots, x_{i|\mathcal{M}_{1,n}|}^n$ according to the distribution $P_{X|V}^n(\cdot|v_i^n)$. We denote the generated code by $\mathcal{C}_{1,n}$. The empirical distribution of the codeword is given by

$$\begin{aligned} P_{\tilde{V}^n \tilde{X}^n}(v^n, x^n) &= \sum_{i \in \mathcal{M}_{2,n}} \sum_{j \in \mathcal{M}_{1,n}} \frac{1}{|\mathcal{M}_{2,n}| |\mathcal{M}_{1,n}|} \mathbf{1}[v_i^n = v^n, x_{ij}^n = x^n], \\ P_{\tilde{V}^n}(v^n) &= \sum_{i \in \mathcal{M}_{2,n}} \frac{1}{|\mathcal{M}_{2,n}|} \mathbf{1}[v_i^n = v^n], \\ P_{\tilde{X}^n|\tilde{V}^n}(x^n|v^n) &= \frac{P_{\tilde{V}^n \tilde{X}^n}(v^n, x^n)}{P_{\tilde{V}^n}(v^n)}, \end{aligned}$$

and the output distribution is given by

$$P_{\tilde{Z}^n}(z^n) = \sum_{v^n, x^n} P_{\tilde{V}^n \tilde{X}^n}(v^n, x^n) P_{Z|X}^n(z^n|x^n).$$

For this construction, we have the following lemma.

Lemma 14: For every $n \geq 1$, we have

$$\begin{aligned} \mathbb{E}_{\mathcal{C}_{1,n} \mathcal{C}_{2,n}} [D(P_{\tilde{Z}^n} \| P_Z^n)] &\leq \frac{1}{\theta |\mathcal{M}_{1,n}|^\theta} e^{n\psi(\theta|P_{Z|X}, P_{X|V}, P_V)} \\ &\quad + \frac{1}{\theta' |\mathcal{M}_{2,n}|^{\theta'}} e^{n\psi(\theta'|P_{Z|V}, P_V)} \end{aligned} \quad (26)$$

⁸Actually, slightly weaker statement, i.e., $D(P_{\tilde{Z}^n} \| P_Z^n)$ in Eq. (23) is replaced by $\frac{1}{n} D(P_{\tilde{Z}^n} \| P_Z^n)$, was proved in [12]. The present statement can be derived from the result in [11].

for $0 < \theta, \theta' \leq 1$, where $\mathbb{E}_{\mathcal{C}_{1,n} \mathcal{C}_{2,n}}[\cdot]$ means taking the average over the randomly generated codes $\mathcal{C}_{1,n}$ and $\mathcal{C}_{2,n}$, the function $\psi(\theta|P_{Z|X}, P_{X|V}, P_V)$ is defined as

$$\begin{aligned} \psi(\theta|P_{Z|X}, P_{X|V}, P_V) &= \log \sum_v P_V(v) \sum_z \left(\sum_x P_{X|V}(x|v) P_{Z|X}(z|x)^{1+\theta} \right) P_{Z|V}(z|v)^{-\theta}, \end{aligned} \quad (27)$$

and $\psi(\theta|\cdot, \cdot)$ is defined in Eq. (25).

Proof: See Appendix A. ■

Corollary 15: If $R_1 > I(X; Z|V)$ and $R_2 > I(V; Z)$, there exists a sequence of map $\varphi_n : \mathcal{M}_{1,n} \times \mathcal{M}_{2,n} \rightarrow \mathcal{X}^n$ such that

$$\lim_{n \rightarrow \infty} D(P_{\tilde{Z}^n} \| P_Z^n) = 0, \quad (28)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_{1,n}| \leq R_1, \quad (29)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_{2,n}| \leq R_2. \quad (30)$$

Proof: See Appendix B. ■

From Corollary 15, we find that the channel resolvability coding scheme proposed in this section can achieve the rate shown in Eq. (22), i.e., $I(X; Z) = I(V; Z) + I(X; Z|V)$. Splitting the randomness into two part does not have any meaning in the channel resolvability coding, but as we will find in Section IV-A, this coding scheme does have meaning when we send the confidential message.

IV. PROOFS OF MAIN RESULTS

A. Proof of Direct Part of Theorem 2

We prove the direct part of Theorem 2 by using the result in Section III. The direct part of the theorem follows from the following Lemma 16 and Lemma 17.

Lemma 16: Let $\mathcal{R}^{(in)}$ be a closed convex set consisting of those quadruples (R_d, R_0, R_1, R_s) for which there exist auxiliary random variables (U, V) such that $U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z)$ and

$$R_0 \leq I(U; Z), \quad (31)$$

$$R_1 + R_s \leq I(V; Y|U), \quad (32)$$

$$R_0 + R_1 + R_s \leq I(V; Y), \quad (33)$$

$$R_1 \geq I(V; Z|U), \quad (34)$$

$$R_d \geq I(X; Z|V). \quad (35)$$

Then $\mathcal{R}^{(in)} \subset \mathcal{R}$.

We note the following observation. From the definition of the problem, if

$$(R_d - r_d, R_0 + r_0, R_1 - r_0 - r_s + r_d, R_s + r_s) \in \mathcal{R}$$

for some $r_d, r_0, r_s \geq 0$, then we also have $(R_d, R_0, R_1, R_s) \in \mathcal{R}$. By using this argument, we have the following.

Lemma 17: We have $\mathcal{R}^* \subset \mathcal{R}$.

Proof: See Appendix C. ■

We now prove Lemma 16. For a while, we consider the case with $n = 1$ and omit the superscript and subscript to simplify the notation. For each common message $k \in \mathcal{K}$, we randomly generate codeword u_k according to distribution P_U . We denote such a code \mathcal{C}_0 . For each k and for each $(\ell, s) \in \mathcal{L} \times \mathcal{S}$, we randomly generate codeword $v_{k\ell s}$ according to distribution $P_{V|U}(\cdot|u_k)$. We denote such a code \mathcal{C}_2 . For each (k, ℓ, s) and for each $a \in \mathcal{A}$, we randomly generate codeword $x_{k\ell s a}$ according to distribution $P_{X|V}(\cdot|v_{k\ell s})$. We denote such a code \mathcal{C}_1 .

For real numbers $\alpha_0, \alpha_1, \alpha_2 \geq 0$ specified later, let

$$\begin{aligned}\mathcal{T}_0 &= \{(u, z) : P_{Z|U}(z|u) \geq e^{\alpha_0} P_Z(z)\}, \\ \mathcal{T}_1 &= \{(u, v, y) : P_{Y|V}(y|v) \geq e^{\alpha_1} P_{Y|U}(y|u)\}, \\ \mathcal{T}_2 &= \{(v, y) : P_{Y|V}(y|v) \geq e^{\alpha_2} P_Y(y)\},\end{aligned}$$

and let $\mathcal{T} = \mathcal{T}_1 \cap (\mathcal{U} \times \mathcal{T}_2)$. Eve's decoding region is defined by

$$\mathcal{D}_k = \{z : (u_k, z) \in \mathcal{T}_0, (u_{\hat{k}}, z) \notin \mathcal{T}_0 \ \forall \hat{k} \neq k\},$$

i.e., $\phi(z) = k$ if $z \in \mathcal{D}_k$. Bob only decode (k, ℓ, s) and he does not decode dummy randomness $a \in \mathcal{A}$. Bob's decoding region is defined by

$$\begin{aligned}\mathcal{D}_{k\ell s} &= \{y : (u_k, v_{k\ell s}, y) \in \mathcal{T}, \\ &\quad (u_{\hat{k}}, v_{\hat{k}\hat{\ell}\hat{s}}, y) \notin \mathcal{T} \ \forall (\hat{k}, \hat{\ell}, \hat{s}) \neq (k, \ell, s)\},\end{aligned}$$

i.e., $g(y) = (k, \ell, s)$ if $y \in \mathcal{D}_{k\ell s}$.

By the above code construction, we have the following.

Lemma 18: We have

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_0 \mathcal{C}_1 \mathcal{C}_2}[P_{err}(f, g)] &\leq P_{UVY}(\mathcal{T}_1^c) + P_{VY}(\mathcal{T}_2^c) \\ &\quad + |\mathcal{L}||\mathcal{S}|e^{-\alpha_1} + |\mathcal{K}||\mathcal{L}||\mathcal{S}|e^{-\alpha_2},\end{aligned}\quad (36)$$

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_0 \mathcal{C}_1 \mathcal{C}_2}[P_{err}(f, \phi)] &\leq P_{UZ}(\mathcal{T}_0^c) + |\mathcal{K}|e^{-\alpha_0}\end{aligned}\quad (37)$$

and

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_0 \mathcal{C}_1 \mathcal{C}_2}[D(f)] &\leq \frac{1}{\theta|\mathcal{A}|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}, P_V)} \\ &\quad + \frac{1}{\theta'|\mathcal{L}|^{\theta'}} e^{\psi(\theta'|P_{Z|V}, P_{V|U}, P_U)}\end{aligned}\quad (38)$$

for $0 < \theta, \theta' \leq 1$, where the functions $\psi(\theta|\cdot, \cdot, \cdot)$ is defined by Eq. (27).

Proof: See Appendix D. \blacksquare

We apply Lemma 18 for asymptotic case. For $(R_d, R_0, R_1, R_s) \in \mathcal{R}^{(in)}$ and arbitrary small $\delta > 0$, we set $|\mathcal{K}_n| = \lfloor e^{n(R_0-2\delta)} \rfloor$, $|\mathcal{L}_n| = \lfloor e^{n(R_1+2\delta)} \rfloor$, $|\mathcal{S}_n| = \lfloor e^{n(R_s-4\delta)} \rfloor$, $|\mathcal{A}_n| = \lfloor e^{n(R_d+2\delta)} \rfloor$, $\alpha_0 = I(U; Z) - \delta$, $\alpha_1 = I(V; Y|U) - \delta$, $\alpha_2 = I(V; Y) - \delta$. Then,

$$\begin{aligned}|\mathcal{L}_n||\mathcal{S}_n|e^{-\alpha_1 n} &\leq e^{-n(I(V; Y|U) - R_1 - R_s + \delta)}, \\ |\mathcal{K}_n||\mathcal{L}_n||\mathcal{S}_n|e^{-\alpha_2 n} &\leq e^{-n(I(V; Y) - R_0 - R_1 - R_s + 3\delta)}, \\ |\mathcal{K}_n|e^{-\alpha_0 n} &\leq e^{-n(I(U; Z) - R_0 + \delta)}\end{aligned}$$

⁹If $z \notin \mathcal{D}_k$ for every $k \in \mathcal{K}$, we set $\phi(z) = 1$, which is not important in our analysis of error probability. A similar remark is also applied for Bob's decoder.

converge to 0 asymptotically. Furthermore, by the law of large numbers, $P_{UVY}^n(\mathcal{T}_{1,n}^c)$, $P_{VY}^n(\mathcal{T}_{2,n}^c)$, and $P_{UZ}^n(\mathcal{T}_{0,n}^c)$ also converge to 0 asymptotically.

Since $\psi'(0|P_{Z|X}, P_{X|V}, P_V) = I(X; Z|V)$, there exists $\theta_0 > 0$ such that

$$\frac{\psi(\theta_0|P_{Z|X}, P_{X|V}, P_V)}{\theta_0} \leq I(X; Z|V) + \delta \leq R_d + \delta,$$

which implies

$$-\frac{\theta_0}{n} \log |\mathcal{A}_n| + \psi(\theta_0|P_{Z|X}, P_{X|V}, P_V) \leq -\delta.$$

Thus,

$$\frac{1}{\theta_0|\mathcal{A}_n|^{\theta_0}} e^{n\psi(\theta_0|P_{Z|X}, P_{X|V}, P_V)}$$

exponentially converges to 0 asymptotically. Similarly, since $\psi'(0|P_{Z|V}, P_{V|U}, P_U) = I(V; Z|U)$, there exists $\theta'_0 > 0$ such that

$$\frac{\psi(\theta'_0|P_{X|V}, P_{V|U}, P_U)}{\theta'_0} \leq I(V; Z|U) + \delta \leq R_1 + \delta,$$

which implies

$$-\frac{\theta'_0}{n} \log |\mathcal{L}_n| + \psi(\theta'_0|P_{X|V}, P_{V|U}, P_U) \leq -\delta.$$

Thus,

$$\frac{1}{\theta'_0|\mathcal{L}_n|^{\theta'_0}} e^{n\psi(\theta'_0|P_{X|V}, P_{V|U}, P_U)}$$

exponentially converges 0 asymptotically. This completes the proof of Lemma 16. \blacksquare

B. Proof of Converse Part of Theorem 2

Suppose that $(R_d, R_0, R_1, R_s) \in \mathcal{R}$. Then, for arbitrary $\gamma > 0$, there exists n such that

$$\begin{aligned}n(R_0 - \gamma) &\leq \log |\mathcal{K}_n|, \\ n(R_0 + R_1 + R_s - \gamma) &\leq \log |\mathcal{K}_n||\mathcal{L}_n||\mathcal{S}_n|, \\ n(R_s - \gamma) &\leq \log |\mathcal{S}_n|, \\ n(R_1 + R_d + \gamma) &\geq \log |\mathcal{L}_n||\mathcal{A}_n|, \\ n(R_d + \gamma) &\geq \log |\mathcal{A}_n|.\end{aligned}$$

By combining these inequalities with the following Lemma 19 and Lemma 20, we have the converse part of the theorem. The statement about the range sizes of U and V can be proved exactly in the same manner as [6, Theorem 17.13]. It should be noted that Eqs. (9)–(11) are derived in the same manner as [6, Theorem 17.13] and the construction of the auxiliary random variables are also the same. Eqs. (12) and (13) are additionally proved in this paper by using the fact that the encoder is deterministic given the dummy randomness.

Lemma 19: There exists $\varepsilon_n \rightarrow 0$ such that

$$\begin{aligned}
\log |\mathcal{K}_n| &\leq I(K_n; Y^n) + n\varepsilon_n, \\
\log |\mathcal{K}_n| &\leq I(K_n; Z^n) + n\varepsilon_n, \\
\log |\mathcal{K}_n| |\mathcal{L}_n| |\mathcal{S}_n| &\leq I(K_n, L_n, S_n; Y^n) + n\varepsilon_n, \\
\log |\mathcal{K}_n| |\mathcal{L}_n| |\mathcal{S}_n| &\leq I(L_n, S_n; Y^n | K_n) + I(K_n; Z^n) + 2n\varepsilon_n, \\
\log |\mathcal{S}_n| &\leq I(L_n, S_n; Y^n | K_n) - I(L_n, S_n; Z^n | K_n) + 4n\varepsilon_n, \\
\log |\mathcal{L}_n| |\mathcal{A}_n| &\geq I(X^n; Z^n | K_n) - 2n\varepsilon_n, \\
\log |\mathcal{A}_n| &\geq I(X^n; Z^n | K_n, L_n, S_n).
\end{aligned}$$

Proof: By using Fano's inequality, we have

$$\begin{aligned}
\log |\mathcal{K}_n| &= H(K_n) \\
&= I(K_n; Y^n) + H(K_n | Y^n) \\
&\leq I(K_n; Y^n) + n\varepsilon_n,
\end{aligned}$$

and

$$\log |\mathcal{K}_n| \leq I(K_n; Z^n) + n\varepsilon_n.$$

By using Fano's inequality, we also have

$$\begin{aligned}
\log |\mathcal{K}_n| |\mathcal{L}_n| |\mathcal{S}_n| &= H(K_n, L_n, S_n) \\
&\leq I(K_n, L_n, S_n; Y^n) + n\varepsilon_n
\end{aligned}$$

and

$$\begin{aligned}
\log |\mathcal{K}_n| |\mathcal{L}_n| |\mathcal{S}_n| &= H(L_n, S_n | K_n) + H(K_n) \\
&= I(L_n, S_n; Y^n | K_n) + I(K_n; Z^n) + 2n\varepsilon_n.
\end{aligned}$$

By using the security condition and Fano's inequality, we have

$$\begin{aligned}
I(S_n; Z^n | K_n) &= I(S_n, K_n; Z^n) - I(K_n; Z^n) \\
&= I(S_n; Z^n) + I(K_n; Z^n | S_n) - I(K_n; Z^n) \\
&\leq I(S_n; Z^n) + H(K_n | Z^n) \\
&\leq 2n\varepsilon_n.
\end{aligned} \tag{39}$$

By using Fano's inequality and by using Eq. (39), we have

$$\begin{aligned}
\log |\mathcal{S}_n| &= H(S_n | K_n) \\
&\leq I(S_n; Y^n | K_n) + n\varepsilon_n \\
&= I(L_n, S_n; Y^n | K_n) - I(L_n; Y^n | S_n, K_n) + n\varepsilon_n \\
&\leq I(L_n, S_n; Y^n | K_n) - H(L_n | S_n, K_n) + 2n\varepsilon_n \\
&\leq I(L_n, S_n; Y^n | K_n) - I(S_n; Z^n | K_n) \\
&\quad - H(L_n | S_n, K_n) + 4n\varepsilon_n \\
&\leq I(L_n, S_n; Y^n | K_n) - I(L_n, S_n; Z^n | K_n) + 4n\varepsilon_n.
\end{aligned}$$

By noting that f_n is a deterministic function and by using Eq. (39), we have

$$\begin{aligned}
\log |\mathcal{L}_n| |\mathcal{A}_n| &\geq H(X^n | K_n, S_n) \\
&\geq I(X^n; Z^n | K_n, S_n) \\
&= I(X^n, S_n; Z^n | K_n) - I(S_n; Z^n | K_n) \\
&\geq I(X^n; Z^n | K_n) - 2n\varepsilon_n.
\end{aligned}$$

Finally, by noting that f_n is a deterministic function, we have

$$\begin{aligned}
\log |\mathcal{A}_n| &\geq H(X^n | K_n, L_n, S_n) \\
&\geq I(X^n; Z^n | K_n, L_n, S_n).
\end{aligned}$$

■

Lemma 20: For fixed n , let T be the random variable that is uniformly distributed over $\{1, \dots, n\}$ and independent of the other random variables. Define the following auxiliary random variables

$$\begin{aligned}
U_t &= (K_n, Y_1^{t-1}, Z_{t+1}^n), \\
V_t &= (L_n, S_n, U_t), \\
U &= (U_T, T), \\
V &= (V_T, T).
\end{aligned}$$

Then, we have

$$I(K_n; Y^n) \leq nI(U; Y), \tag{40}$$

$$I(K_n; Z^n) \leq nI(U; Z), \tag{41}$$

$$I(K_n, L_n, S_n; Y^n) \leq n[I(V; Y | U) + I(U; Y)], \tag{42}$$

$$I(L_n, S_n; Y^n | K_n) + I(K_n; Z^n) \leq n[I(V; Y | U) + I(U; Z)], \tag{43}$$

$$I(L_n, S_n; Y^n | K_n) - I(L_n, S_n; Z^n | K_n) \leq n[I(V; Y | U) - I(V; Z | U)], \tag{44}$$

$$I(X^n; Z^n | K_n) \geq nI(X; Z | U), \tag{45}$$

$$I(X^n; Z^n | K_n, L_n, S_n) \geq nI(X; Z | V) \tag{46}$$

and

$$U \leftrightarrow V \leftrightarrow X \leftrightarrow (Y, Z). \tag{47}$$

Proof: Since the proof of Eq. (47) is well known [6], we only prove the other inequalities.

a) *Proof of Eq. (40):*

$$\begin{aligned}
& I(K_n; Y^n) \\
&= \sum_{t=1}^n I(K_n; Y_t | Y_1^{t-1}) \\
&\leq \sum_{t=1}^n I(K_n, Y_1^{t-1}, Z_{t+1}^n; Y_t) \\
&= \sum_{t=1}^n I(U_t; Y_t) \\
&= nI(U_T; Y_T | T) \\
&= nI(U_T, T; Y_T) \\
&= nI(U; Y).
\end{aligned}$$

b) *Proof of Eq. (41):*

$$\begin{aligned}
& I(K_n; Z^n) \\
&= \sum_{t=1}^n I(K_n; Z_t | Z_{t+1}^n) \\
&\leq \sum_{t=1}^n I(K_n, Y_1^{t-1}, Z_{t+1}^n; Z_t) \\
&= \sum_{t=1}^n I(U_t; Z_t) \\
&= nI(U_T; Z_T | T) \\
&= nI(U_T, T; Z_T) \\
&= nI(U; Z).
\end{aligned}$$

c) *Proof of Eq. (42):*

$$\begin{aligned}
& I(K_n, L_n, S_n; Y^n) \\
&= \sum_{t=1}^n I(K_n, L_n, S_n; Y_t | Y_1^{t-1}) \\
&\leq \sum_{t=1}^n I(K_n, L_n, S_n, Y_1^{t-1}, Z_{t+1}^n; Y_t) \\
&= \sum_{t=1}^n I(U_t, V_t; Y_t) \\
&= nI(U_T, V_T; Y_T | T) \\
&= nI(U, V; Y) \\
&= n[I(V; Y|U) + I(U; Y)].
\end{aligned}$$

d) *Proof of Eq. (43):*

$$\begin{aligned}
& I(L_n, S_n; Y^n | K_n) + I(K_n; Z^n) \\
&= \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}) + I(K_n; Z_t | Z_{t+1}^n)] \\
&\leq \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}, Z_{t+1}^n) \\
&\quad + I(Z_{t+1}^n; Y_t | K_n, Y_1^{t-1}) + I(K_n; Z_T | Z_{t+1}^n)] \\
&\stackrel{(a)}{=} \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_t^{t-1}, Z_{t+1}^n) \\
&\quad + I(Y_1^{t-1}; Z_t | K_n, Z_{t+1}^n) + I(K_n; Z_t | Z_{t+1}^n)]
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}, Z_{t+1}^n) \\
&\quad + I(K_n, Y_1^{t-1}, Z_{t+1}^n; Z_t)] \\
&= \sum_{t=1}^n [I(V_t; Y_t | U_t) + I(U_t; Z_t)] \\
&= n[I(V_T; Y_T | U_T, T) + I(U_T; Z_T | T)] \\
&= n[I(V; Y | U) + I(U; Z)],
\end{aligned}$$

where we used Csiszár's sum identity [24] in (a).

e) *Proof of Eq. (44):*

$$\begin{aligned}
& I(L_n, S_n; Y^n | K_n) - I(L_n, S_n; Z^n | K_n) \\
&= \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}) - I(L_n, S_n; Z_t | K_n, Z_{t+1}^n)] \\
&\stackrel{(a)}{=} \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}) \\
&\quad + I(Z_{t+1}^n; Y_t | K_n, L_n, S_n, Y_1^{t-1}) \\
&\quad - I(Y_1^{t-1}; Z_t | K_n, L_n, S_n, Z_{t+1}^n) \\
&\quad - I(L_n, S_n; Z_t | K_n, Z_{t+1}^n)] \\
&= \sum_{t=1}^n [I(L_n, S_n, Z_{t+1}^n; Y_t | K_n, Y_1^{t-1}) \\
&\quad - I(L_n, S_n, Y_1^{t-1}; Z_t | K_n, Z_{t+1}^n)] \\
&\stackrel{(b)}{=} \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}, Z_{t+1}^n) \\
&\quad + I(Z_{t+1}^n; Y_t | K_n, Y_1^{t-1}) - I(Y_1^{t-1}; Z_t | K_n, Z_{t+1}^n) \\
&\quad - I(L_n, S_n; Z_t | K_n, Y_1^{t-1}, Z_{t+1}^n)] \\
&= \sum_{t=1}^n [I(L_n, S_n; Y_t | K_n, Y_1^{t-1}, Z_{t+1}^n) \\
&\quad - I(L_n, S_n; Z_t | K_n, Y_1^{t-1}, Z_{t+1}^n)] \\
&= \sum_{t=1}^n [I(V_t; Y_t | U_t) - I(V_t; Z_t | U_t)] \\
&= n[I(V_T; Y_T | U_T, T) - I(V_T; Z_T | U_T, T)] \\
&= n[I(V; Y | U) - I(V; Z | U)],
\end{aligned}$$

where (a) and (b) follow from Csiszár's sum identity [24].

f) *Proof of Eq. (45):*

$$\begin{aligned}
& I(X^n; Z^n | K_n) \\
&= \sum_{t=1}^n [H(Z_t | K_n, Z_{t+1}^n) - H(Z_t | K_n, X^n, Z_{t+1}^n)] \\
&\stackrel{(a)}{\geq} \sum_{t=1}^n [H(Z_t | K_n, Y_1^{t-1}, Z_{t+1}^n) - H(Z_t | K_n, X_t, Y_1^{t-1}, Z_{t+1}^n)] \\
&= \sum_{t=1}^n I(X_t; Z_t | K_n, Y_1^{t-1}, Z_{t+1}^n) \\
&= \sum_{t=1}^n I(X_t; Z_t | U_t) \\
&= nI(X_T; Z_T | U_T, T) \\
&= nI(X; Z | U),
\end{aligned}$$

where (a) follows from the fact that $(K_n, X_1^{t-1}, X_{t+1}^n, Y_1^{t-1}, Z_{t+1})$, X_t , and Z_t form Markov chain.

g) *Proof of Eq. (46):*

$$\begin{aligned}
& I(X^n; Z^n | K_n, L_n, S_n) \\
&= \sum_{t=1}^n [H(Z_t | K_n, L_n, S_n, Z_{t+1}^n) \\
&\quad - H(Z_t | K_n, L_n, S_n, X_t^n, Z_{t+1}^n)] \\
&\stackrel{(a)}{\geq} \sum_{t=1}^n [H(Z_t | K_n, L_n, S_n, Y_1^{t-1}, Z_{t+1}^n) \\
&\quad - H(Z_t | K_n, L_n, S_n, X_t, Y_1^{t-1}, Z_{t+1}^n)] \\
&= \sum_{t=1}^n I(X_t; Z_t | K_n, L_n, S_n, Y_1^{t-1}, Z_{t+1}^n) \\
&= \sum_{t=1}^n I(X_t; Z_t | V_t) \\
&= nI(X_T; Z_T | V_T, T) \\
&= nI(X; Z | V),
\end{aligned}$$

where (a) follows from the fact that $(K_n, L_n, S_n, X_1^{t-1}, X_{t+1}^n, Y_1^{t-1}, Z_{t+1}^n)$, X_t , and Z_t form Markov chain. ■

V. CONCLUSION

In this paper, we investigated the trade-off between the rate of the random number, the rates of common, private, and confidential messages.

As a by-product of our result, Lemma 14 can be also applied to the three receiver wire-tap channel, and the lower bound of secrecy capacity obtained in [7, Corollary 1] with strong security can be proved.

ACKNOWLEDGMENT

This research was initiated by a discussion with Prof. Ryutaroh Matsumoto about the deterministic encoding result in [5]. The authors would like to thank him for bringing the authors' attention to the randomness constrained stochastic encoding problem. This research is partly supported by Grand-in-Aid for Young Scientists(B):2376033700, Grant-in-Aid for Scientific Research(B):2336017202, and Grant-in-Aid for Scientific Research(A):2324607101.

APPENDIX

A. Proof of Lemma 14

For simplicity of notation, we only prove the statement for $n = 1$, and the subscript n is omitted in the proof. The statement for $n \geq 2$ can be proved by regarding the n th product distribution as one distribution and by noting that

$$\psi(\theta | P_{Z|X}^n, P_{X|V}^n, P_V^n) = n\psi(\theta | P_{Z|X}, P_{X|V}, P_V)$$

and

$$\psi(\theta' | P_{Z|V}^n, P_V^n) = n\psi(\theta' | P_{Z|V}, P_V)$$

hold.

We first note the following observations. By taking average over the randomly generated codes \mathcal{C}_1 and \mathcal{C}_2 , we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} [P_{\tilde{V}}(v) P_{\tilde{X}|\tilde{V}}(x|v)] \\
&= \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} \left[\sum_{i \in \mathcal{M}_2} \sum_{j \in \mathcal{M}_1} \frac{1}{|\mathcal{M}_1| |\mathcal{M}_2|} \mathbf{1}[v_i = v, x_{ij} = x] \right] \\
&= P_{VX}(v, x)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} [P_{\tilde{V}}(v)] &= \mathbb{E}_{\mathcal{C}_2} [P_{\tilde{V}}(v)] \\
&= \mathbb{E}_{\mathcal{C}_2} \left[\sum_{i \in \mathcal{M}_2} \frac{1}{|\mathcal{M}_2|} \mathbf{1}[v_i = v] \right] \\
&= P_V(v).
\end{aligned} \tag{48}$$

Furthermore, for fixed \mathcal{C}_2 , by taking the average over the randomly generated code \mathcal{C}_1 , we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_1} [P_{\tilde{X}|\tilde{V}}(x|v)] \\
&= \mathbb{E}_{\mathcal{C}_1} \left[\frac{\sum_{i \in \mathcal{M}_2} \sum_{j \in \mathcal{M}_1} \frac{1}{|\mathcal{M}_1| |\mathcal{M}_2|} \mathbf{1}[v_i = v, x_{ij} = x]}{P_{\tilde{V}}(v)} \right] \\
&= \frac{\sum_{i \in \mathcal{M}_2} \sum_{j \in \mathcal{M}_1} \frac{1}{|\mathcal{M}_1| |\mathcal{M}_2|} \mathbf{1}[v_i = v] P_{X|V}(x|v)}{P_{\tilde{V}}(v)} \\
&= P_{X|V}(x|v)
\end{aligned} \tag{49}$$

Let $P_{Z'}$ be the output distribution when the input distribution is $P_{\tilde{V}}(v) P_{X|V}(x|v)$. Then, from Eq. (49), we have

$$\mathbb{E}_{\mathcal{C}_1} [P_{\tilde{Z}}(z)] = P_{Z'}(z)$$

for every $z \in \mathcal{Z}$. Thus, we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} [D(P_{\tilde{Z}} \| P_Z)] \\
&= \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} \left[\sum_z P_{\tilde{Z}}(z) \log \frac{P_{\tilde{Z}}(z)}{P_Z(z)} \right] \\
&= \mathbb{E}_{\mathcal{C}_2} \left[\mathbb{E}_{\mathcal{C}_1} \left[\sum_z P_{\tilde{Z}}(z) \log \frac{P_{\tilde{Z}}(z)}{P_{Z'}(z)} \right] \right. \\
&\quad \left. + \mathbb{E}_{\mathcal{C}_1} \left[\sum_z P_{\tilde{Z}}(z) \log \frac{P_{Z'}(z)}{P_Z(z)} \right] \right] \\
&= \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} \left[\sum_z P_{\tilde{Z}}(z) \log \frac{P_{\tilde{Z}}(z)}{P_{Z'}(z)} \right] \\
&\quad + \mathbb{E}_{\mathcal{C}_2} \left[\sum_z P_{Z'}(z) \log \frac{P_{Z'}(z)}{P_Z(z)} \right] \\
&= \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} [D(P_{\tilde{Z}} \| P_{Z'})] + \mathbb{E}_{\mathcal{C}_2} [D(P_{Z'} \| P_Z)]. \tag{50}
\end{aligned}$$

We bound each term of Eq. (50) by using Proposition 13. By the monotonicity of the divergence, we have

$$\begin{aligned}
D(P_{\tilde{Z}} \| P_{Z'}) &\leq D(P_{I\tilde{Z}} \| P_{I Z'}) \\
&= \sum_{i \in \mathcal{M}_2} \frac{1}{|\mathcal{M}_2|} D(P_{\tilde{Z}|I}(\cdot|i) \| P_{Z|V}(\cdot|v_i)),
\end{aligned}$$

where

$$P_{I\tilde{Z}}(i, z) = \frac{1}{|\mathcal{M}_2|} \sum_{j \in \mathcal{M}_1} \frac{1}{|\mathcal{M}_1|} P_{Z|X}(z|x_{ij})$$

and

$$P_{IZ'}(i, z) = \frac{1}{|\mathcal{M}_2|} P_{Z|V}(z|v_i).$$

Thus, by taking average over \mathcal{C}_1 and by using Proposition 13 for input distribution $P_{X|V}(\cdot|v_i)$ instead of P_X , we have

$$\begin{aligned} & \mathbb{E}_{\mathcal{C}_1} [D(P_{\bar{Z}} \| P_{Z'})] \\ & \leq \sum_{i \in \mathcal{M}_2} \frac{1}{|\mathcal{M}_2|} \mathbb{E}_{\mathcal{C}_1} [D(P_{\bar{Z}|I}(\cdot|i) \| P_{Z|V}(\cdot|v_i))] \\ & \leq \sum_{i \in \mathcal{M}_2} \frac{1}{|\mathcal{M}_2|} \frac{1}{|\mathcal{M}_1|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}(\cdot|v_i))}. \end{aligned}$$

By taking average over \mathcal{C}_2 and by noting Eq. (48), we have

$$\begin{aligned} & \mathbb{E}_{\mathcal{C}_1 \mathcal{C}_2} [D(P_{\bar{Z}} \| P_{Z'})] \\ & \leq \sum_v P_V(v) \frac{1}{\theta |\mathcal{M}_1|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}(\cdot|v))} \\ & = \frac{1}{\theta |\mathcal{M}_1|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}, P_V)}. \end{aligned}$$

On the other hand, by using Proposition 13 for input distribution P_V and channel $P_{Z|V}$, we have

$$\mathbb{E}_{\mathcal{C}_2} [D(P_{Z'} \| P_Z)] \leq \frac{1}{\theta' |\mathcal{M}_2|^{\theta'}} e^{\psi(\theta'|P_{Z|V}, P_V)}.$$

■

B. Proof of Corollary 15

We can choose $\gamma > 0$ such that $R_1 \geq I(X; Z|V) + 2\gamma$ and $R_2 \geq I(V; Z) + 2\gamma$. Let $|\mathcal{M}_{1,n}| = \lfloor e^{nR_1} \rfloor$, $|\mathcal{M}_{2,n}| = \lfloor e^{nR_2} \rfloor$. Since $\psi'(0|P_{Z|X}, P_{X|V}, P_V) = I(X; Z|V)$, there exists $\theta_0 > 0$ such that

$$\frac{\psi(\theta_0|P_{Z|X}, P_{X|V}, P_V)}{\theta_0} \leq I(X; Z|V) + \gamma \leq R_1 - \gamma,$$

which implies

$$-\frac{\theta_0}{n} \log |\mathcal{M}_{1,n}| + \psi(\theta_0|P_{Z|X}, P_{X|V}, P_V) \leq -\gamma < 0.$$

Thus, the first term of Eq. (26) converges to 0 asymptotically. Similarly, we can show that the second term of Eq. (26) converges to 0 asymptotically. Thus, we have the assertion of the corollary. ■

C. Proof of Lemma 17

Although the lemma can be systematically proved by using the Fourier-Motzkin elimination, we explicitly find (r_d, r_0, r_s) satisfying

$$(R_d - r_d, R_0 + r_0, R_1 - r_0 - r_s + r_d, R_s + r_s) \in \mathcal{R}^{(in)} \quad (51)$$

for given $(R_d, R_0, R_1, R_s) \in \mathcal{R}^*$ as follows.

If $R_1 + R_s \leq I(V; Y|U)$ and $R_1 < I(V; Z|U)$, we set

$$r_d := I(V; Z|U) - R_1$$

and $(r_0, r_s) := (0, 0)$. Then, Eqs. (31), and (34) are obviously satisfied. Eq. (32) can be confirmed as

$$\begin{aligned} R_1 + r_d + R_s & \leq I(V; Z|U) + I(V; Y|U) - I(V; Z|U) \\ & = I(V; Y|U), \end{aligned}$$

Eq (33) can be confirmed as

$$\begin{aligned} R_0 + R_1 + r_d + R_s & \leq I(U; Y) + I(V; Y|U) \\ & = I(V; Y), \end{aligned}$$

and Eq. (35) can be confirmed as

$$\begin{aligned} R_d - r_d & = R_d + R_1 - I(V; Z|U) \\ & \geq I(X; Z|U) - I(V; Z|U) \\ & = I(X; Z|V). \end{aligned}$$

Thus, Eq. (51) holds.

If $R_1 + R_s \leq I(V; Y|U)$ and $R_1 \geq I(V; Z|U)$, we set $(r_d, r_1, r_s) := (0, 0, 0)$. Then, Eq. (51) obviously holds.

If $R_1 + R_s > I(V; Y|U)$, we set $r_d := 0$ and

$$\begin{aligned} r_s & := I(V; Y|U) - I(V; Z|U) - R_s, \\ r_0 & := R_1 + R_s - I(V; Y|U) \\ & = (R_1 - r_s) + (R_s + r_s) - I(V; Y|U) \\ & = R_1 - r_s - I(V; Z|U). \end{aligned} \quad (52)$$

Then, Eqs. (32), (33), and (35) are obviously satisfied. Eq. (31) can be confirmed as

$$\begin{aligned} R_0 + r_0 & = R_0 + R_1 + R_s - I(V; Y|U) \\ & \leq I(U; Z), \end{aligned}$$

and Eq. (34) can be confirmed from Eq. (52). Thus, Eq. (51) is satisfied.

D. Proof of Lemma 18

a) *Proof of Eq. (36):* We first note the following observations. By taking the average over randomly generated codes, we have

$$\begin{aligned} & \mathbb{E}_{\mathcal{C}_0 \mathcal{C}_1 \mathcal{C}_2} [P_{err}(f, g)] \\ & = \mathbb{E}_{\mathcal{C}_0 \mathcal{C}_1 \mathcal{C}_2} \left[\sum_{k, \ell, s, a} \frac{1}{|\mathcal{K}| |\mathcal{L}| |\mathcal{S}| |\mathcal{A}|} P_{Y|X}(\mathcal{D}_{k\ell s}^c | x_{k\ell s a}) \right] \\ & = \mathbb{E}_{\mathcal{C}_0 \mathcal{C}_2} \left[\sum_{k, \ell, s, a} \frac{1}{|\mathcal{K}| |\mathcal{L}| |\mathcal{S}| |\mathcal{A}|} \mathbb{E}_{\mathcal{C}_1} [P_{Y|X}(\mathcal{D}_{k\ell s}^c | x_{k\ell s a})] \right] \\ & = \mathbb{E}_{\mathcal{C}_0 \mathcal{C}_2} \left[\sum_{k, \ell, s} \frac{1}{|\mathcal{K}| |\mathcal{L}| |\mathcal{S}|} P_{Y|V}(\mathcal{D}_{k\ell s}^c | v_{k\ell s}) \right]. \end{aligned} \quad (53)$$

Let $\mathcal{T}_{uv} = \{y : (u, v, y) \in \mathcal{T}\}$. Then, we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_0\mathcal{C}_2} \left[\sum_{k,\ell,s} \frac{1}{|\mathcal{K}||\mathcal{L}||\mathcal{S}|} P_{Y|V}(\mathcal{D}_{k\ell s}^c | v_{k\ell s}) \right] \\
& \leq \mathbb{E}_{\mathcal{C}_0\mathcal{C}_1} \left[\sum_{k,\ell,s} \frac{1}{|\mathcal{K}||\mathcal{L}||\mathcal{S}|} \{P_{Y|V}(\mathcal{T}_{u_k v_{k\ell s}}^c | v_{k\ell s}) \right. \\
& \quad \left. + \sum_{(\hat{k}, \hat{\ell}, \hat{s}) \neq (k, \ell, s)} P_{Y|V}(\mathcal{T}_{u_{\hat{k}} v_{\hat{k}\hat{\ell}\hat{s}}} | v_{k\ell s}) \} \right] \\
& = \mathbb{E}_{\mathcal{C}_0\mathcal{C}_2} \left[\sum_{k,\ell,s} \frac{1}{|\mathcal{K}||\mathcal{L}||\mathcal{S}|} \{P_{Y|V}(\mathcal{T}_{u_k v_{k\ell s}}^c | v_{k\ell s}) \right. \\
& \quad \left. + \sum_{(\hat{\ell}, \hat{s}) \neq (\ell, s)} P_{Y|V}(\mathcal{T}_{u_k v_{k\hat{\ell}\hat{s}}} | v_{k\ell s}) \right. \\
& \quad \left. + \sum_{k \neq \hat{k}} \sum_{\hat{\ell}, \hat{s}} P_{Y|V}(\mathcal{T}_{u_{\hat{k}} v_{\hat{k}\hat{\ell}\hat{s}}} | v_{k\ell s}) \} \right] \\
& \leq \sum_{k,\ell,s} \frac{1}{|\mathcal{K}||\mathcal{L}||\mathcal{S}|} \{P_{UVY}(\mathcal{T}^c) \\
& \quad + |\mathcal{L}||\mathcal{S}| \sum_{u,v} P_{UV}(u, v) P_{Y|U}(\mathcal{T}_{uv} | u) \\
& \quad + |\mathcal{K}||\mathcal{L}||\mathcal{S}| \sum_{u,v} P_{UV}(u, v) P_Y(\mathcal{T}_{uv}) \} \\
& \leq P_{UVY}(\mathcal{T}^c) + |\mathcal{L}||\mathcal{S}| e^{-\alpha_1} + |\mathcal{K}||\mathcal{L}||\mathcal{S}| e^{-\alpha_2},
\end{aligned}$$

where we used

$$\begin{aligned}
P_{Y|U}(y|u) & \leq P_{Y|V}(y|v) e^{-\alpha_1}, \\
P_Y(y) & \leq P_{Y|V}(y|v) e^{-\alpha_2}
\end{aligned}$$

for $y \in \mathcal{T}_{uv}$ in the last inequality.

b) *Proof of Eq. (37):* In a similar manner as Eq. (53), we have

$$\mathbb{E}_{\mathcal{C}_0\mathcal{C}_1\mathcal{C}_2} [P_{err}(f, \phi)] = \mathbb{E}_{\mathcal{C}_0} \left[\sum_k \frac{1}{|\mathcal{K}|} P_{Z|U}(\mathcal{D}_k^c | u_k) \right],$$

which is just a random coding error probability of channel $P_{Z|U}$. Thus, by the standard arguments of the information spectrum approach [25], we have Eq. (37).

c) *Proof of Eq. (38):* By the monotonicity of the divergence, we have

$$\begin{aligned}
& D(P_{S\tilde{Z}} \| P_S \times P_{\tilde{Z}}) \\
& \leq D(P_{KS\tilde{Z}} \| P_S \times P_{K\tilde{Z}}) \\
& = \sum_k \frac{1}{|\mathcal{K}|} D(P_{S\tilde{Z}|K}(\cdot | k) \| P_S \times P_{\tilde{Z}|K}(\cdot | k)) \\
& = \sum_{k,s} \frac{1}{|\mathcal{K}||\mathcal{S}|} D(P_{\tilde{Z}|KS}(\cdot | k, s) \| P_{\tilde{Z}|K}(\cdot | k)).
\end{aligned}$$

Note that the relation

$$\begin{aligned}
& \sum_s \frac{1}{|\mathcal{S}|} D(P_{\tilde{Z}|KS}(\cdot | k, s) \| P_{\tilde{Z}|K}(\cdot | k)) \\
& \quad + D(P_{\tilde{Z}|K}(\cdot | k) \| P_{Z|U}(\cdot | u_k)) \\
& = \sum_s \frac{1}{|\mathcal{S}|} D(P_{\tilde{Z}|KS}(\cdot | k, s) \| P_{Z|U}(\cdot | u_k))
\end{aligned}$$

holds for each $k \in \mathcal{K}$. Thus, by using Lemma 14 for $P_{V|U}(\cdot | u_k)$ instead of P_V , we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_1\mathcal{C}_2} [D(f)] \\
& \leq \sum_k \frac{1}{|\mathcal{K}|} \left[\frac{1}{\theta|\mathcal{A}|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}, P_{V|U}(\cdot | u_k))} \right. \\
& \quad \left. + \frac{1}{\theta'|\mathcal{L}|^{\theta'}} e^{\psi(\theta'|P_{Z|V}, P_{V|U}(\cdot | u_k))} \right].
\end{aligned}$$

By taking average over \mathcal{C}_0 and by noting

$$\mathbb{E}_{\mathcal{C}_0} \left[\sum_k \frac{1}{|\mathcal{K}|} \mathbf{1}[u_k = u] \right] = P_U(u),$$

we have

$$\begin{aligned}
& \mathbb{E}_{\mathcal{C}_0\mathcal{C}_1\mathcal{C}_2} [D(f)] \\
& \leq \sum_u P_U(u) \left[\frac{1}{\theta|\mathcal{A}|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}, P_{V|U}(\cdot | u))} \right. \\
& \quad \left. + \frac{1}{\theta'|\mathcal{L}|^{\theta'}} e^{\psi(\theta'|P_{Z|V}, P_{V|U}(\cdot | u))} \right] \\
& = \frac{1}{\theta|\mathcal{A}|^\theta} e^{\psi(\theta|P_{Z|X}, P_{X|V}, P_V)} \\
& \quad + \frac{1}{\theta'|\mathcal{L}|^{\theta'}} e^{\psi(\theta'|P_{Z|V}, P_{V|U}, P_U)}.
\end{aligned}$$

■

E. Proof of Corollary 12

By noting that U, V, X , and (Y, Z) form Markov chain, which implies $I(V; Y|X, U) = 0$ and $I(V; Z|X, U) = 0$, we have

$$\begin{aligned}
& I(V; Y|U) - I(V; Z|U) \\
& = I(V, X; Y|U) - I(V, X; Z|U) \\
& \quad - [I(X; Y|U, V) - I(X; Z|U, V)] \\
& = I(X; Y|U) - I(X; Z|U) \\
& \quad - [I(X; Y|U, V) - I(X; Z|U, V)].
\end{aligned}$$

Since $P_{Y|X}$ is more capable than $P_{Z|X}$, we have

$$I(X; Y|U = u, V = v) - I(X; Z|U = u, V = v) \geq 0$$

for every (u, v) , which implies

$$I(V; Y|U) - I(V; Z|U) \leq I(X; Y|U) - I(X; Z|U).$$

Thus, the auxiliary random variable V is not needed. ■

REFERENCES

- [1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [2] Y. Liang, H. V. Poor, and S. S. (Shitz), *Information Theoretic Security*. Now Publishers, 2009.
- [3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339–348, May 1979.
- [4] Y. Steinberg and S. Verdú, "Channel simulation and coding with side-information," *IEEE Trans. Inform. Theory*, vol. 40, no. 3, pp. 634–646, May 1994.
- [5] Y. Oohama and S. Watanabe, "Capacity results for relay channels with confidential messages," 2009, arXiv:1009.5829.
- [6] I. Csiszár and J. Körner, *Information Theory, Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.
- [7] Y.-K. Chia and A. E. Gamal, "Three-receiver broadcast channel with common and confidential messages," *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp. 2748–2765, May 2012.
- [8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.
- [9] U. Maurer, "The strong secret key rate of discrete random triples," in *Communication and Cryptography—Two Sides of One Tapestry*. Kluwer Academic Publishers, 1994, pp. 271–285.
- [10] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, 1996.
- [11] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans. Inform. Theory*, vol. 57, no. 6, pp. 3989–4001, June 2011, arXiv:0904.0308.
- [12] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
- [13] N. Cai, A. Winter, and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *Problems of Information Transmission*, vol. 40, no. 4, pp. 26–47, 2004.
- [14] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 44–55, January 2005, arXiv:quant-ph/0304127.
- [15] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.
- [16] M. R. Bloch, "Achieving secrecy: Capacity vs. resolvability," in *Proceedings of IEEE International Symposium on Information Theory*, Saint Petersburg, Russia, August 2011, pp. 632–636.
- [17] Y. Oohama, "Converse coding theorem for identification via general degraded broadcast channels," in *IEEE Int. Symp. Inform. Theory*, Yokohama, Japan, July 2003.
- [18] M. Hayashi and R. Matsumoto, "Secure multiplex coding with dependent and non-uniform multiple messages," arXiv:1202.1332.
- [19] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to attain the ordinary channel capacity securely in wiretap channels," in *Proc. IEEE Information Theory Workshop on Theory and Practice in Information Theoretic Security*, 2005, pp. 13–18, arXiv:cs/0509047.
- [20] M. Bloch and J. Kliever, "On secure communication with constrained randomization," arXiv:1202.5529v1.
- [21] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. 23, no. 1, pp. 60–64, January 1977.
- [22] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4207–4214, September 2010.
- [23] O. Ozel and S. Ulukus, "Wiretap channels: Roles of rate splitting and channel prefixing," in *Proc. IEEE Int. Symp. Inf. Theory 2011*, Saintpeterburg, Russia, 2011, pp. 627–631, arXiv:1110.4613.
- [24] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, 2011.
- [25] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, 2003.